



Courtier, M., Croxford, A., & Atherton, K. (2017). An iterative design of experiments based data collection approach for ultrasonic guided waves. In *43rd Annual Review of Progress in Quantitative Nondestructive Evaluation: (17–22 July 2016, Atlanta, Georgia, USA)* (Vol. 36). [030012] (AIP Conference Proceedings; Vol. 1806). American Institute of Physics (AIP). <https://doi.org/10.1063/1.4974580>

Peer reviewed version

Link to published version (if available):  
[10.1063/1.4974580](https://doi.org/10.1063/1.4974580)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via AIP at <http://aip.scitation.org/doi/abs/10.1063/1.4974580>. Please refer to any applicable terms of use of the publisher.

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

# An Iterative Design of Experiments Based Data Collection Approach for Ultrasonic Guided Waves

Mark R. Courtier<sup>1,a)</sup>, Anthony J. Croxford<sup>1</sup> and Kathryn Atherton<sup>2</sup>

<sup>1</sup>*Department of Mechanical Engineering, University of Bristol, Queen's Building, University Walk, Bristol, BS8 1TR, UK*

<sup>2</sup>*Airbus Operations Ltd., Pegasus House, Aerospace Avenue, Filton, Bristol, BS34 7PA, UK*

<sup>a)</sup>Corresponding author: m.r.courtier@bristol.ac.uk

**Abstract.** To validate a model of guided wave propagation from acoustic emission on real aerospace structures a large amount of experimental data must be collected. It is desirable to take measurements over large areas at a small spacing between measurement points and this is very time consuming. To reduce the experimental effort, the HilomotDoE algorithm has been adapted to provide a black box method of selecting the data points to be collected. It chooses points which do not fit well to local hinging plane models. For example this includes the drop in amplitude of a guided wave as it crosses a stringer. The algorithm has been tested on a simulated structure and experimentally on a real structure. On the simulated structure, which had a relatively low feature density, using the algorithm reduces the number of data points required to 17.3% of the number of points in a full raster scan to reach a sufficiently good representation of the full raster scan. On the experimental structure, the algorithm did not reduce the number of points collected significantly and this was due to the much higher feature density in this structure. This highlights that the algorithm will only reduce the number of collection points required where appropriate. This algorithm has potential in reducing the number of data points collected in other NDT applications where measurement time is large and parts of the response are predictable.

## INTRODUCTION

For the development and deployment of Acoustic Emission (AE) structural health monitoring systems it can be beneficial to model the guided wave propagation within the structure. This gives a better understanding of detection range and propagation paths of waves from AE sources at different positions within the structure [1, 2]. This can be used to inform system design and interpret results from AE tests. To validate models, experimental data must be collected and, due to the size of the application structures and density of measurement points often required, this can be very time consuming. This paper describes and demonstrates a technique which can reduce the number of measurement points and therefore total experiment time.

In previous work [3], a model of the guided wave propagation from AE sources in complex aerospace structures has been created. Experimental data to validate the model has been collected using an Electromagnetic Acoustic Transducer (EMAT) as a proxy for an AE source. An EMAT has the advantages that it can be easily be moved from one measurement position to the next, it has very consistent coupling and its sensitivity is reasonably omnidirectional. It has the disadvantage of a low transduction efficiency when compared to piezoelectric based transducers. This means the combination of a narrowband chirp signal [4] and averaging multiple measurements is required to get a received signal with a sufficiently high Signal to Noise Ratio (SNR). The areas over which measurements are required are relatively large, in the order of 1 m<sup>2</sup>, and the EMAT is positioned manually. The pitch between measurements can be required to be as small as half a wavelength. This combination of measurement time and positioning time means each measurement is relatively time consuming. Depending on the background noise level, measurements can take 1 to 3 minutes. This situation has led to investigations on how to reduce the time to conduct the whole data collection, resulting in the development of an iterative Design of Experiments (DoE) based data collection approach.

The most common experimental design is grid based, also known as a raster scan. This is a type of geometrical experimental design. Here measurements are taken at a certain pitch which is constant across the measurement space.

This method comprehensively covers the measurement space but is the most time consuming permutation. Alternative geometrical experimental designs exist including Fractional Factorial and Latin Hypercube [5]. These types of experimental design reduce the number of sampling points and therefore time required but do not consider the process output in choosing these points. Given the features of interest in the target application are in small regions, these techniques are unlikely to sample at a sufficient resolution in these regions.

Another common group of experimental design techniques is optimal experimental designs. Here a model of the output is assumed and the sample points are chosen to minimise a certain type of variance of this output [5]. These type of experimental designs reduce the number of sampling points required but require a model type to be assumed. This model type is normally a simple polynomial model, which is unlikely to fit the output response in this application, but can be a more specific model. Given that the results of this experimental work will be used to test a model of guided wave propagation, it is preferable not to have to predict another model of the response. If this was done there is a risk of missing unexpected results and conformational bias in the chosen sampling points.

For the experiments described above much of the structure is a plate of constant thickness. Here the wave propagation is influenced by beam spread and attenuation, both of which are well understood [2]. Therefore measurements in the centre of plates are generally not showing interesting information and these regions can be sampled less densely. Geometrical features are likely to have a more significant affect on the wave propagation [6] so measurements in these regions are of more interest. Denser sampling is therefore required in these regions. To achieve this without making assumptions about the response requires the use of another type of DoE approach; an experimental design with an active learning step.

The algorithm chosen here is the *Hierarchical local model tree for Design of Experiments* (HilomotDoE) [5, 7]. The operation of the algorithm will be described in detail in the following section but it consists of a set of local models which are fitted to subsets of the dataset. These local models are adaptive in size and resolution and the local model which worst fits its dataset is improved on each iteration by collecting more data. This means the algorithm will collect data points where changes are present which should correspond to the regions where the guided wave amplitude is changing. It will collect less data points in regions where the amplitude is constant or changing linearly. This is the desired collection behaviour of collecting fewer points but the most relevant points.

The next section will describe the details of the algorithm and then the following sections will demonstrate the algorithm on simulated and experimentally collected data.

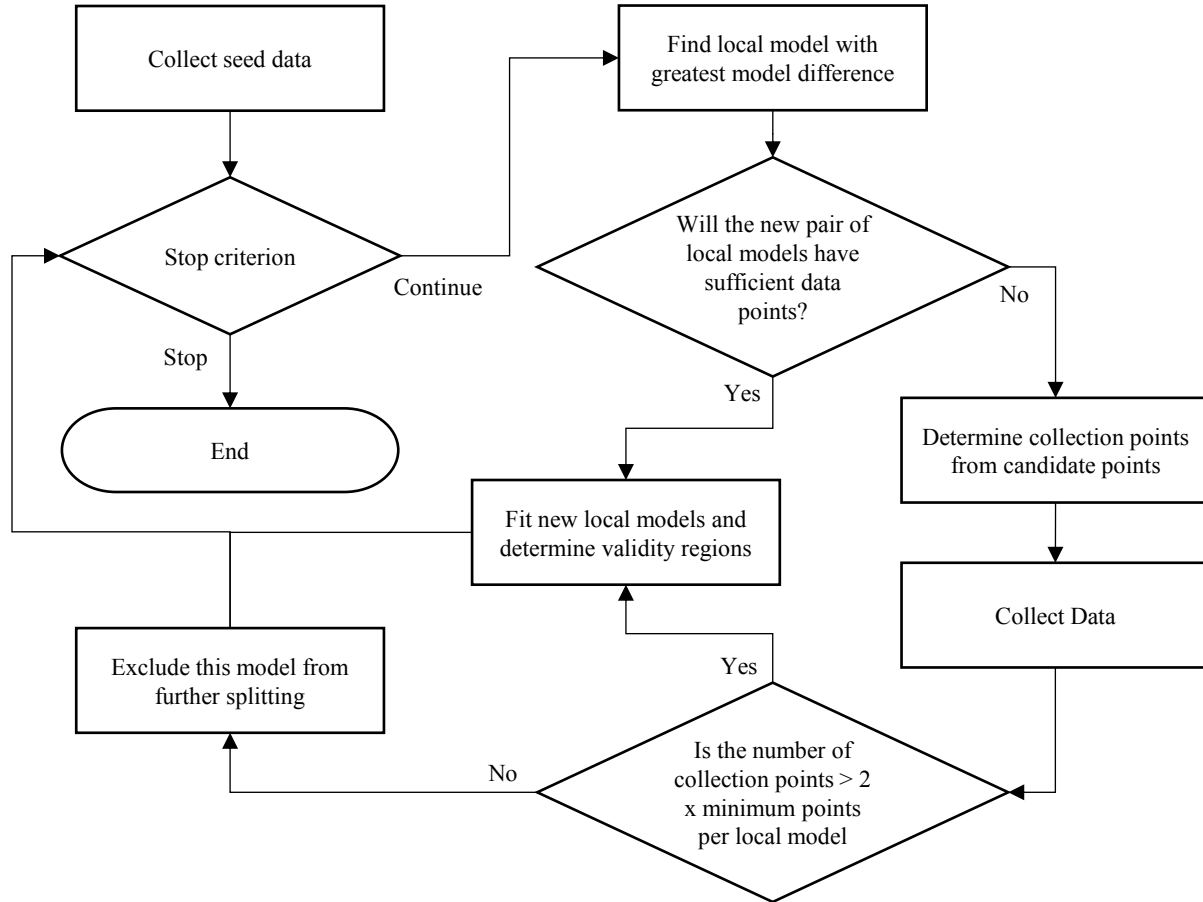
## ALGORITHM OPERATION

The algorithm operates in the following way:

- First a seed dataset is collected which sparsely covers the region to be measured.
- A model, which can be of any type, is fitted to the dataset to give a global model. The same type of model is used throughout the algorithm. The Root Mean Squared (RMS) difference between the global model and the data set is calculated.
- If this global model difference is too large, the model is refined by splitting the model in half. The optimum split is found so that the two new local models created have the minimum difference between the model and the experimental results. The partitioning strategy is based upon the one described by Nelles in [8]. A minimum number of points per local model is defined and if the new local models to be created will not have sufficient points, more measurements are made prior to the split being calculated.
- A new global model is formed from the local models and the validity functions created by the partitioning strategy. The global model difference is recalculated and if refinement is still required then the local model with the worst error will be split in half.
- The iterations continue until a stop criterion is met. This could be convergence of the model difference or a maximum number of samples to take.

## Algorithm Implementation

The implementation of the algorithm used to collect data for this report will now be explained including the deviations from the HilomotDoE algorithm explained above. The operation of the data collection algorithm is shown in Fig. 1.



**FIGURE 1.** Flow chart showing the operation of the data collection algorithm.

### Collect seed data

A minimum resolution is defined by the size of the smallest feature of interest. It is important to have at least one sample point on each feature otherwise the algorithm may not choose that region to collect more points and therefore the feature would be missed. The minimum spacing between points is used to define an initial set of sample points at which measurements are taken. The points are determined by the modified pseudo-Monte Carlo sampling algorithm explained below.

### Stop criterion

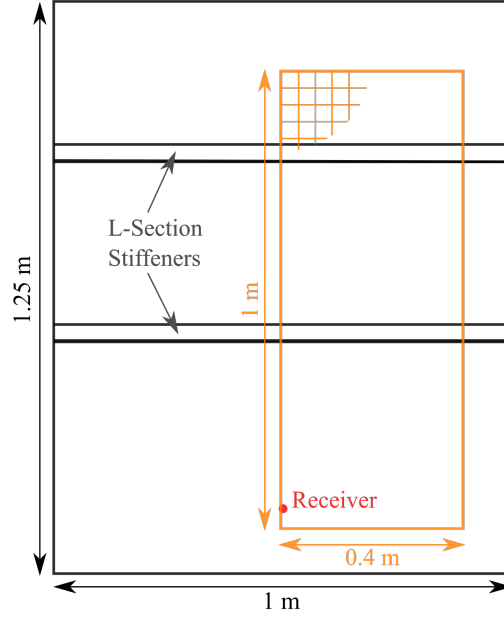
The decision whether to continue can be based upon many factors including the time available for experimentation, a threshold on the value of model difference or the convergence of the model difference. In this paper a threshold will be used but further work is required in this area.

### Find local model with greatest model difference

To determine which local model is to be split, the local model with the greatest RMS model difference is found. This is the criteria used to determine which local model least well represents the underlying process.

### Determine collection points from candidate points

When a model region is to be split into two local models, new data collection points are required if there will not be sufficient points in the two new models to satisfy the defined minimum number of points per model. It is necessary to have a minimum number of points per model to ensure that the model does not become over-fitted which would damage the validity of the model difference calculation. The new collection points are chosen from a list of candidate points which are determined by a user defined resolution which will be half the minimum wavelength of the frequency range of interest in most cases. The candidate point chosen is that with the greatest nearest neighbour distance to the already collected points. This is repeated until the necessary number of new



**FIGURE 2.** Diagram of simulated experiment including the receiver position, source locations and L-section stiffeners.

candidate points have been defined. This method is used so that the points picked are well spaced and is inspired by the pseudo-Monte Carlo sampling algorithm in [5] but with a more constrained set of candidate points. If multiple points have an equal greatest nearest neighbour distance then one of these points is chosen at random. If there are not sufficient available candidate points to satisfy the minimum number of points per model then the available points are collected and the local model is excluded from further splitting.

#### **Collect data**

The ultrasonic time traces for each transducer in the array is collected and saved for each measurement position. The variable of interest is then calculated from the time trace. In this case this is the maximum amplitude of the first arrival.

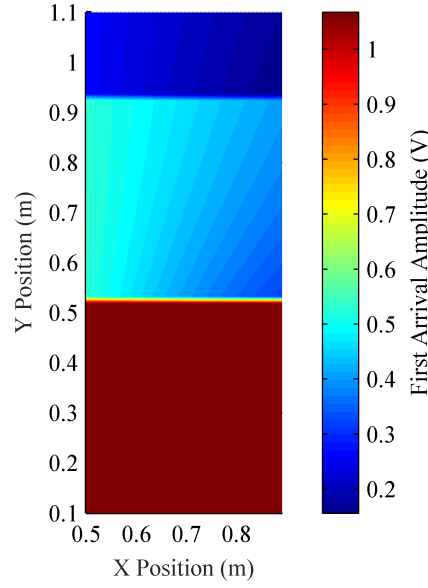
#### **Fit new local models and determine validity regions**

The best pair of local models is found by minimising the RMS model difference for the two new local models. First the validity functions for each new region are calculated. The split is constrained to pass through the centre of the points to ensure that they are split in half and therefore each new local model satisfies the minimum number of points per model. Then the best fit model is found for the data. The model type used here is a hinging hyperplane [9]. This is a pair of planes which meet at a line like a piece of paper with a single fold. Hinging planes are used because they do not fit well over discontinuities, for example the amplitude drop as the guided waves travel over a feature, and therefore force data collection in this region. An iterative search is performed to find the optimum pair of validity functions. This and the corresponding local models are included in and the local model they replace is removed from the global model.

The exception to this procedure is the first iteration. Here the global model consists of one local model which is valid everywhere. Because there is no split between models, the local model is a plane fitted to the data and not a hinging hyperplane.

## **EXAMPLE ON SIMULATED DATA**

First the algorithm will be applied to a simulated dataset to demonstrate its performance. The advantage of using a simulated dataset is that the full raster scan can be generated in a short period of time, negating the issue which the algorithm aims to address. The structure modelled is a 3 mm aluminium plate which has had 2 L-section stiffeners bolted to it. The stiffeners are also made of 3 mm aluminium and are 38 mm wide and tall. It is attached with M5 bolts which have a spacing of 25 mm between them. The structure is shown in Fig. 2. The wave propagation in the



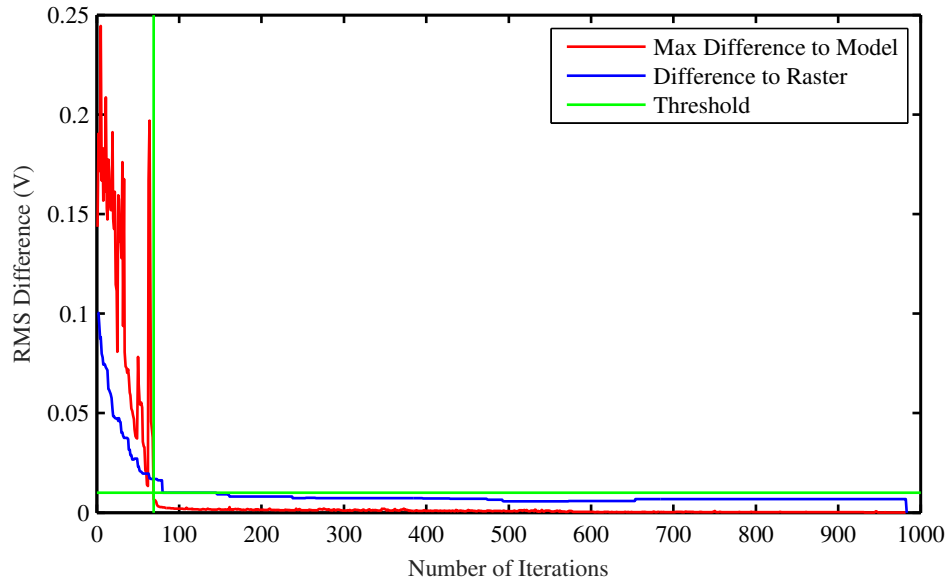
**FIGURE 3.** The amplitude of the first  $S_0$  arrival on the plate with 2 L-section stiffeners after the effect of beam spread has been removed.

structure is modelled using the linear time-invariant systems approach for the direct ray paths from the AE source to the receiving transducer. Each component is modelled in the frequency domain including an empirical model for the transmission across the L-section. The full modelling methodology is described in [3]. The excitation signal was a 5 cycle 200 kHz Hanning windowed tone burst. The raster scan image of the  $S_0$  first arrival amplitude is shown in Fig. 3. Here the affect of beam spread has been removed because this is a predictable asymptotic relationship.

The algorithm has been applied to the simulated structure and it picks locations from the raster scan iteratively, as it would request measurements to be made during a real experiment. Components within the algorithm are random and because the full raster scan is available, the algorithm will be run over the same dataset multiple times to check the convergence of the model difference is consistent. This is the metric that informs the stop criterion so reliable convergence of this value is required for the algorithm to be successfully applied. The target number of points per local model was set to 12 because this was found to perform well.

### Convergence of Model Difference

Figures 4 and 5 show how the difference between the sparse dataset collected by the algorithm at that iteration to the values predicted by the model change as the number of iterations the algorithm is applied increases. This is for one run of the algorithm and multiple runs of the algorithm respectively. The difference to model is calculated at each point in the full raster scan and where data has not yet been collected in the sparse dataset, the value is found from a linear interpolation, based upon a triangulation. From this the RMS difference is found for each local model. The maximum RMS difference across the local models is then found and this is the value shown in Fig. 4. If all possible points in a local model are collected it is excluded from this calculation. The algorithm has been run with no stop criterion and therefore eventually collects all available points. It can be seen that the maximum RMS difference to model decreases to a very small value prior to 100 iterations. The value decreases in quite a chaotic way as refinement of the model discovers larger differences than were previously present in the sparse dataset. No significant difference is present after 100 iterations. The RMS difference between the sparse dataset and the raster scan decreases in a similar way. This information is not known to the algorithm as it requires a full raster scan to be measured but demonstrates how well the sparse dataset represents the full raster scan. The difference to raster converges at a similar number of iterations to the difference to model showing that this is a good metric to base the stop criterion upon. The difference to raster converges to a value slightly larger than zero until the very end of the algorithm operation. This is because of errors in the interpolation from the sparse dataset which are only corrected once all points are collected. This difference is



**FIGURE 4.** How the difference of model to sparse dataset and raster scan to sparse dataset changes as the algorithm progresses. This is the result for a simulation of a single experiment.

acceptably small when considering the amplitude of the signals.

The algorithm has been applied to this simulated dataset 924 times. The maximum model difference has been found for each run and the mean and 95% quantiles of this value can then be found. This is shown in Fig. 5. It can be seen that the behaviour of the difference to model and difference to raster are similar to that of the single run but mean convergence occurs at a slightly higher number of iterations. Both variable converge at a similar number of iterations confirming that the difference to model is a good metric to base a stop criterion upon. It should be noted that each run of the algorithm stops at a different number of iterations so the greater variability in the quantiles at the largest number of iterations is caused by a smaller sample size.

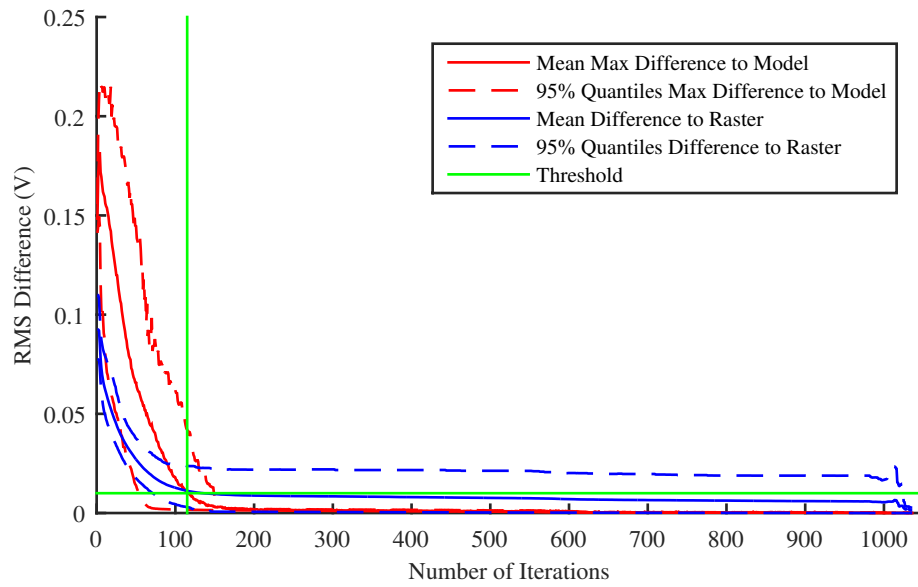
### Threshold as a Stop Criterion

One proposed stop criterion for the data collection is a threshold on the maximum difference to model. A threshold of 0.01 V was set which is approximately 100 times smaller than the maximum first arrival amplitude. This was deemed to be a sufficiently small error. In the single run of the algorithm on the simulated dataset, shown in Fig. 4, the iteration where the threshold was crossed was 69 which corresponds to 600 data points collected out of a possible 5040 which is 11.9% of the total number of points. Across the 924 runs of the algorithm, shown in Fig. 5, the mean number of data points required to reach the threshold was 870 which is 17.3% of the total number of points. This is a significant reduction in the number of points that need to be collected and therefore a significant reduction in total experiment time.

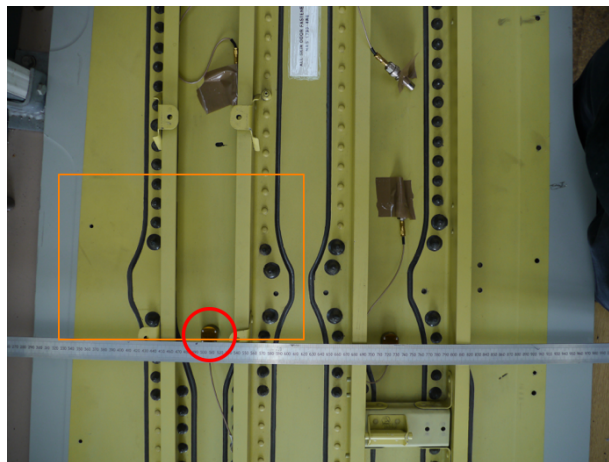
Other potential stop criteria exist and it is believed it would be worth investigating a stop criterion based upon the convergence of the maximum difference to model. This would potentially remove the requirement to define a threshold which currently has to be done intuitively from an estimate of the signal amplitude.

### EXAMPLE ON EXPERIMENTAL DATA

The algorithm will now be applied to an experimentally collected dataset to demonstrate its performance on real signals which contain noise. The dataset is a raster scan collected on a section of A320 wing skin. A roving EMAT transducer was used as a repeatable representation of an AE source. It was excited with a 250 kHz 5 cycle Hanning windowed tone burst. Due to time constraints the region was sampled at 20 mm in the x direction and 40 mm in the

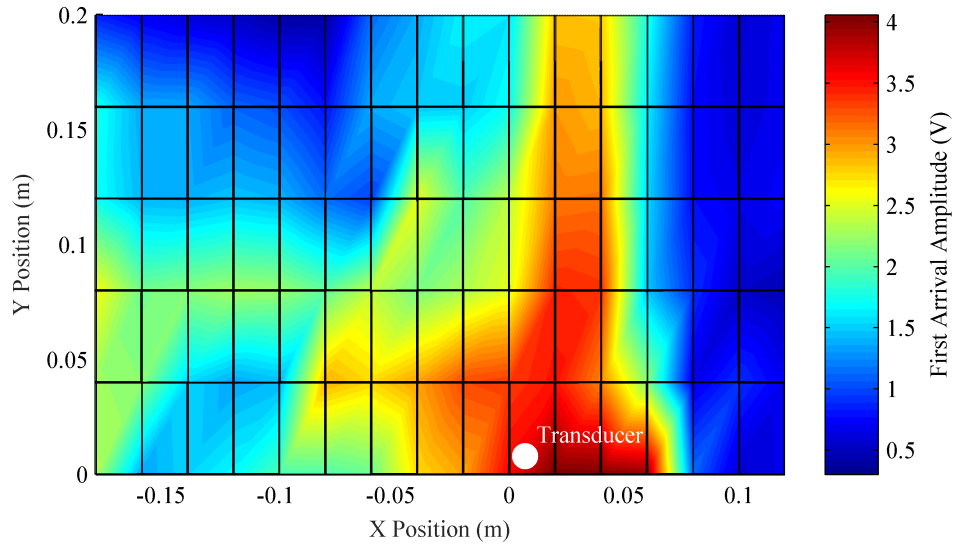


**FIGURE 5.** How the difference of model to sparse dataset and raster scan to sparse dataset changes as the algorithm progresses. The results here are from 924 simulations of the same experiment.

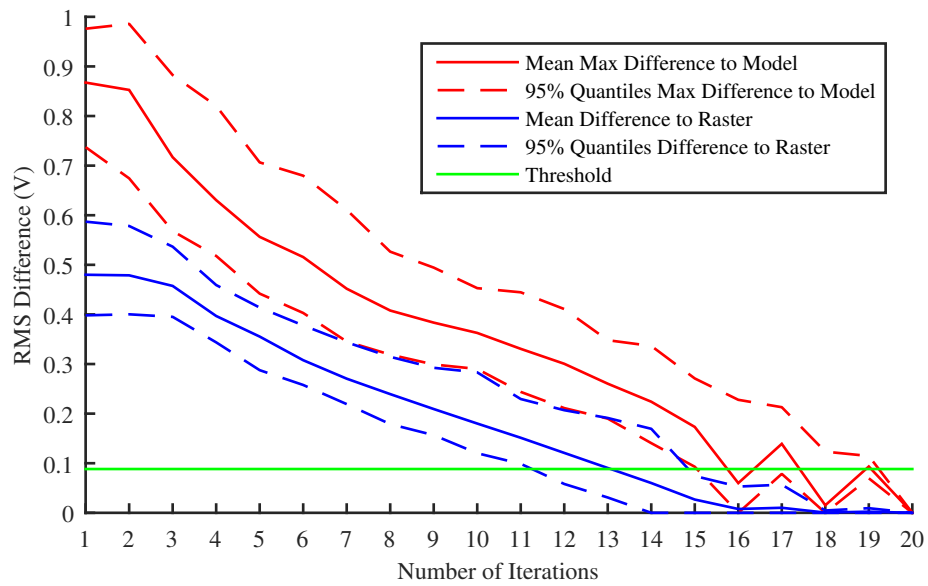


**FIGURE 6.** A diagram of the collection region and receiver position on a section of A320 wing skin.





**FIGURE 7.** The amplitude of the first  $S_0$  arrival on a section of A320 wing skin after the effect of beam spread has been removed.



**FIGURE 8.** How the difference of model to sparse dataset and raster scan to sparse dataset changes as the algorithm progresses. The results here are from 100 runs on the same experimental dataset.

y direction. This is a larger spacing than a half wavelength of the centre frequency of excitation. The wavelength at 250 kHz on this sample is approximately 25 mm. The collection region is shown in Fig. 6 and the amplitude of the first arrival is shown in Fig. 7.

As was done with the simulated plate with 2 L-section stiffeners, the algorithm has been applied to this dataset multiple times to get a clear understanding of how it behaves. The algorithm was applied 100 times to the dataset and the mean and 95% quantiles can be seen in Fig. 8. The target number of points for local models was set to 9. This was found to perform well on this structure with a high feature density. A threshold was defined by the RMS value of the noise in the experimental signals and this was 0.0883 V. The RMS value of the noise is a suitable value for the threshold because the difference to each local model cannot be better than this. Collecting more points after this threshold is reached is not likely to reduce the model difference further.

It can be seen in Fig. 8 that the mean maximum difference to model and mean difference to the full raster scan decrease in a similar pattern as the results from the simulated dataset. Both values decrease at a similar rate. However these values do not converge to a constant value prior to all data points having been collected. The mean maximum difference to model decreases to the threshold value at a similar point to where some of the runs have finished and the mean value is becoming more erratic due to the small number of runs which get to this high a number of iterations. The algorithm would therefore only have stopped in a small number of runs. The values do not converge because of the higher feature density in this sample and that it is sampled less densely. This is a demonstration of where the algorithm will not reduce the number of points to be collected and a full raster scan is required, which is what would have had to be collected without this technique. The results still suggest that the mean maximum difference to model is a suitable metric for a stop criteria.

## CONCLUSIONS

The HilomotDoE algorithm has been applied to reduce the number of data points collected in raster scans to validate a model of AE guided waves. The algorithm collects data points in regions where the first arrival amplitude does not fit plane models well. This is typically the case at features such as stiffeners so more data points are collected here than in parts of the structure which are just plate. This is the desired behaviour because the guided wave propagation is simple to model in the plate regions and the effect at features is of more interest in validating that model.

The algorithm has been applied to both a simulated dataset with 2 L-section stiffeners and a section of A320 wing skin with stringers. In both cases the maximum RMS difference to the model of the response and the RMS difference between the sparse dataset and the full raster scan decrease in a similar way. This suggests that the maximum RMS difference to the model is a suitable metric to base a stop criterion upon and a threshold stop criterion has been demonstrated on the simulated dataset. The feature density on the A320 wing skin means that a full or nearly full raster scan is required so a stop criterion cannot be tested here. With the threshold stop criterion there is an average reduction in the required data points of 82.7% compared to the full raster scan for the simulated dataset.

This approach to data collection has potential to significantly reduce the amount of data points collected and therefore total experimental time in guided wave measurement. The amount of time saved is dependent on the density of features in the structure. It does this in a way which requires no knowledge of the guided wave behaviour in the structure being measured which is an advantage in both simplicity and that it prevent conformational bias in selection of measurement points. Here this algorithm has been applied for a very specific need but it has potential to reduce the number of data points collected in other NDT applications where measurement time is large and parts of the response are predictable.

## FUTURE WORK

The work on this algorithm is ongoing and future work includes:

- Optimising the number of points per local model. This is dependent on the feature density of the structure and maximum resolution of measurement.
- As mentioned previously, there are multiple possible stop criterion and these should be fully investigated.
- Further experimental validation is required including an example which converges prior to a full raster scan being collected.

## ACKNOWLEDGEMENTS

Mark Courtier is enrolled in the Engineering Doctorate Programme in NDE at University of Bristol. He is funded by the Engineering and Physical Sciences Research Council (EPSRC) through the UK Centre for Doctoral Training in Quantitative Non-destructive Evaluation and by Airbus Operations Ltd.

## REFERENCES

1. R. Dalton, P. Cawley, and M. Lowe, IEE Proceedings-Science, Measurement and Technology **148**, 169–177 (2001).
2. J. Scholey, P. Wilcox, M. Wisnom, and M. Friswell, Ultrasonics **49**, 538–548 (2009).
3. M. R. Courtier, A. J. Croxford, and K. Atherton, “Guided wave propagation modelling to aid understanding of acoustic emission system performance on complex aerospace structures,” in *8th European Workshop On Structural Health Monitoring (EWSHM 2016)* (2016).
4. J. E. Michaels, S. J. Lee, A. J. Croxford, and P. D. Wilcox, Ultrasonics **53**, 265–270 (2013).
5. B. Hartmann, J. Moll, O. Nelles, and C.-P. Fritzen, “Hierarchical local model trees for design of experiments in the framework of ultrasonic structural health monitoring,” in *2011 IEEE International Conference on Control Applications (CCA)* (IEEE, 2011), pp. 1163–1170.
6. R. Dalton, P. Cawley, and M. Lowe, Journal of Nondestructive Evaluation **20**, 29–46 (2001).
7. O. Nelles, *Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models* (Springer, 2001).
8. O. Nelles, “Axes-oblique partitioning strategies for local model networks,” in *2006 IEEE Conference on Computer Aided Control System Design, 2006 IEEE International Conference on Control Applications, 2006 IEEE International Symposium on Intelligent Control* (IEEE, 2006), pp. 2378–2383.
9. S. Ernst, “Hinging hyperplane trees for approximation and identification,” in *Decision and Control, 1998. Proceedings of the 37th IEEE Conference on*, Vol. 2 (IEEE, 1998), pp. 1266–1271.